Exact Dynamic Support Tracking with Multiple Measurement Vectors using Compressive MUSIC

Jong Min Kim, Ok Kyun Lee and Jong Chul Ye

arXiv:1110.0378v1 [cs.IT] 3 Oct 2011

Abstract

Dynamic tracking of sparse targets has been one of the important topics in array signal processing. Recently, compressed sensing (CS) approaches have been extensively investigated as a new tool for this problem using partial support information obtained by exploiting temporal redundancy. However, most of these approaches are formulated under single measurement vector compressed sensing (SMV-CS) framework, where the performance guarantees are only in a probabilistic manner. The main contribution of this paper is to allow *deterministic* tracking of time varying supports with multiple measurement vectors (MMV) by exploiting multi-sensor diversity. In particular, we show that a novel compressive MUSIC (CS-MUSIC) algorithm with optimized partial support selection not only allows removal of inaccurate portion of previous support estimation but also enables addition of newly emerged part of unknown support. Numerical results confirm the theory.

Index Terms

Compressed sensing, joint sparsity, time varying signal, compressive MUSIC, optimized partial support selection

Correspondence to:
Jong Chul Ye, Ph.D. Associate Professor
Dept. of Bio and Brain Engineering, KAIST
373-1 Guseong-dong Yuseong-gu, Daejon 305-701, Korea
Email: jong.ye@kaist.ac.kr
Tel: 82-42-350-4320
Fax: 82-42-350-4310

I. INTRODUCTION

Dynamic target tracking problem that addresses the estimation of time varying support of moving target has been one of the important classical topics in array signal processing including radar, communication, and medical imaging applications. For example, in electroencephalography (EEG) or magnetoencephalography (MEG) source localization problems, it has been shown that the position of the dipole moments during epileptic activities varies according to time and we are interested in their spatio-temporal dynamics [1]. Dynamic MRI problem that tracks the motion of hearts also belongs to this class of problem.

Recently, there have been renewed interests for this problem with the help of a modern mathematical tool called compressed sensing [2], [3]. These approaches try to exploit knowledges of partial support information obtained at the previous time point. More specifically, consider the following time varying support estimation problem:

$$\min_{\mathbf{x}(t)} \ \|\mathbf{x}(t)\|_0, \quad \text{subject to} \quad \mathbf{b}(t) = A\mathbf{x}(t), \quad t = 0, 1, \cdots, \tag{1}$$

where $\mathbf{b}(t) \in \mathbb{R}^m$, and $\mathbf{x}(t) \in \mathbb{R}^n$ are noiseless measurement vector, and sparse signal at time t. Assuming that the support is assumed to change slowly, theoretical results [4] have demonstrated that we can reduce the required sampling in compressed sensing reconstruction if we have partially known support from the prior estimation results. For example, Vaswani and Lu proposed modified-CS algorithm [4] which addresses the exact reconstruction of noiseless case with partially known support:

$$\min_{\mathbf{x}(t)} \ \|(\mathbf{x}(t))_{I(t-1)^c}\|_0, \quad \text{subject to} \quad \mathbf{b}(t) = A\mathbf{x}(t), \quad t = 1, 2, \cdots, \tag{2}$$

where $I(t-1)$ is the previously estimated support, and $(\mathbf{x}(t))_{I(t-1)^c}$ denotes a subvector after removing the elements that correspond to the index set $I(t-1)$. Suppose, furthermore, $k = |\text{supp}\mathbf{x}(t)|_0$, $u = |I(t) \setminus I(t-1)|$, and $e = |I(t-1) \setminus I(t)|$. Then, if the restricted isometry constant (RIP) for the sensing matrix A satisfies

$$\delta_{k+e+u} < 1, \tag{3}$$

then the solution $\mathbf{x}(t)$ of Eq. (2) is the unique solution [4]. This is much weaker than $0 \leq \delta_{2k} < 1$ for the original SMV-CS problem [5], in case of slowly time varying support with $u \ll k$ and

$e \ll k$. They further showed an l_1 convex relaxation of Eq. (2) can provide the same l_0 solution of Eq. (2), if the following RIP condition is satisfied:

$$2\delta_{2u} + \delta_{3u} + \delta_{k+e-u} + \delta_{k+e}^2 + 2\delta_{k+e+u}^2 < 1, \tag{4}$$

which is again relaxed sampling requirement than that of original CS problem $\delta_{2k} < \sqrt{2} - 1$ [5]. Therefore, exploiting the temporal redundancy has significant impact for reducing sampling requirement for dynamic support tracking.

Rather than solving the tracking problem Eq. (1) at each time, batch type approaches such as T-SBL (temporal sparse Bayesian learning) [1] collect the multiple snapshot data (for example, $\{\mathbf{b}(t)\}_{t=1}^N$) and process them together to estimate the dynamic varying support. Note that if the support changes slowly over time, then the resulting collection of problem becomes an multiple measurement vector problem. Accordingly, T-SBL converts the resulting MMV problem into a block-sparse SMV problem, after which each block statistics are modeled using a specific Gaussian form temporal correlation structure. The update rule using the expectation-maximization (EM) method and its accelerated version can be then used to solve the resulting Bayesian problem [1].

However, these approaches for dynamic support tracking is with SMV-CS framework and their performance guarantees is in a probabilistic sense. In practice, there are many situations where we can obtain multiple measurement vector information for time varying objects. For example, in single-input multiple-output (SIMO) multiple access channel (MAC), multiple antenna can observe linear combination of individual codewords multiplied by the unknown channel gain from the individual user [6]. In parallel MR cardiac imaging, multiple coils simultaneous obtain k-space measurements of temporally varying hearts with distinct coil sensitivities. In EEG/MEG source localization problem, the dipole moments can be assumed relatively stationary during a short time window from which multiple snapshot of the sensor measurement can be obtained. All these examples acquire multiple measurement of the unknown signal vectors that share the same support with different weighting through identical sensing matrices.

A fundamental question under this setup is what kind of diversity gain we can obtain over SMV-CS support tracking. To our knowledge, we are not aware of any prior investigation in this regard. One of the main contributions of this paper is to show that a multiple measurement vector (MMV) framework not only extend the SMV counterpart, but also provides a unique advantage of "deterministic" support tracking for slow varying support estimation. Recall that MMV can

measure multiple information of a set of vector that share the same sparsity pattern through the identical sensing matrix. This paper shows that this joint sparsity pays off significantly in dynamic support tracking by relaxing probabilistic guarantee to a deterministic guarantee. The feasibility of the exact support tracking has significant impacts in practice.

The breakthrough is based on our novel compressive multiple signal classification (CS-MUSIC) algorithm in MMV compressed sensing problem [7], in which a part of supports are found probabilistically using the conventional CS, after which the remaining supports are determined deterministically using the generalized MUSIC criterion. In addition, CS-MUSIC allows us to find all k support as long as at least $k - r + 1$ support out of any k-support estimate are correct [8], where r denote the rank of the measurement matrix. This result provides an important clue for deterministic and exact dynamic support tracking under MMV setup, in which the probabilistic compressed sensing support estimation step is replaced by the support estimate from the previous snapshots, after which the CS-MUSIC algorithm eliminates the incorrect portion of previous time point support estimation and then add newly updated support deterministically. This update scheme guarantees the exact support tracking in noiseless case under an appropriate sampling condition. Other contributions of our method include that the support error does not propagate along time due to the self-correction step. Furthermore, using large system model, we can derive conditions with which the proposed algorithm correct track the time varying support even in noisy cases. We believe that with these noticeable advantages of our algorithm we may find many important applications in radar, communication as well as biomedical application.

This paper consist of following. Section II reviews the compressive MUSIC and support correction criterion for MMV setup. In Section III, we derive our main theoretical results on sampling condition for deterministic support tracking. Numerical results are given in Section IV, which is followed by conclusion in Section V.

A. Notations and Mathematical Preliminaries

Throughout the paper, $\mathbf{x}^i$ and $\mathbf{x}_j$ correspond to the i-th row and the j-th column of matrix X, respectively. When S is an index set, X^S, A_S corresponds to a submatrix collecting corresponding rows of X and columns of A, respectively. The following definitions are also used throughout the paper.

Definition 1: [9] The rows (or columns) in $\mathbb{R}^n$ are in general position if any n collection of rows (or columns) are linearly independent.

Definition 2: [10] $\mathrm{Spark}(A)$ denotes the smallest number of linearly dependent columns of a matrix A.

Definition 3 (Restricted Isometry Property (RIP)): A sensing matrix $A \in \mathbb{R}^{m \times n}$ is said to have a k-restricted isometry property (RIP) if there exist left and right RIP constants $0 < \delta_k^L, \delta_k^R < 1$ such that

$$(1 - \delta_k^L)\|\mathbf{x}\|^2 \leq \|A\mathbf{x}\|^2 \leq (1 + \delta_k^R)\|\mathbf{x}\|^2$$

for all $\mathbf{x} \in \mathbb{R}^n$ such that $\|\mathbf{x}\|_0 \leq k$. A single RIP constant $\delta_k = \max\{\delta_k^L, \delta_k^R\}$ is often referred to as the RIP constant.

II. MMV COMPRESSIVE SENSING USING COMPRESSIVE MUSIC: A REVIEW

Let m, n and r be a positive integers ($m < n$) that represents the number of sensor elements, the ambient space dimension, and the number of snapshots, respectively. Suppose that we are given a multiple-measurement vector $B \in \mathbb{R}^{m \times r}$, $X = [\mathbf{x}_1, \cdots, \mathbf{x}_r] \in \mathbb{R}^{n \times r}$, and a sensing matrix $A \in \mathbb{R}^{m \times n}$. A canonical form MMV problem [7] is given by the following optimization problem:

$$\text{minimize} \quad \|X\|_0 \tag{5}$$

$$\text{subject to} \quad B = AX,$$

where $\|X\|_0 = |\mathrm{supp}X| = k$, $\mathrm{supp}X = \{1 \leq i \leq n : \mathbf{x}^i \neq 0\}$, and the measurement matrix B is full rank, i.e. $\mathrm{rank}(B) = r \leq \|X\|_0$.

Recall that every MMV problem can be converted to a canonical form MMV using a singular value decomposition and dimension reduction as described in [7]. Now, We can easily expect that the diversity due to the joint sparsity can improve the recovery performance over SMV compressed sensing. Indeed, Chen and Huo [11], Feng and Bresler [12] and recently Davies and Elder [13] showed that $X \in \mathbb{R}^{n \times r}$ is the unique solution of $AX = B$ if and only if

$$\|X\|_0 < \frac{\mathrm{spark}(A) + \mathrm{rank}(B) - 1}{2} \leq \mathrm{spark}(A) - 1 . \tag{6}$$

Note that we can expect $\mathrm{rank}(B)/2$ gains over SMV thanks to the MMV diversity. Furthermore, Feng and Bresler [12] showed that the noiseless l_0 bound in Eq. (6) is achievable using MUSIC algorithm as long as $r = \mathrm{rank}(B) = k$. More specifically, suppose that the columns of a sensing

matrix $A \in \mathbb{R}^{m \times n}$ are in general position. Then, according to [12], [14], for any $j \in \{1, \cdots, n\}$, $j \in \mathrm{supp}X$ if and only if

$$Q^* \mathbf{a}_j = 0, \tag{7}$$

where $Q \in \mathbb{R}^{m \times (m-r)}$ consists of orthonormal columns such that $Q^* B = 0$ so that $R(Q)^\perp = R(B)$, which is often called "noise subspace". Using the compressive sensing terminology, Eq. (7) implies that the recoverable sparsity level by MUSIC (with a probability 1 for the noiseless measurement case) is given by

$$\|X\|_0 < m = \mathrm{spark}(A) - 1, \tag{8}$$

where the last equality comes from the definition of the spark. Therefore, the l_0 bound (6) can be achieved by MUSIC bound in (8) when $r = k$ [12].

However, for any $r < k$, the MUSIC condition (7) does not hold. This is a major drawback of MUSIC compared to CS algorithms that allow perfect reconstruction with a extremely large probability by increasing the sensor elements m. One the other hand, even thought the conventional CS algorithms for MMV such as simultaneous OMP (S-OMP), p-thresholding [15], [16] have good recovery performance when $r \ll k$, but they exhibit performance saturation as r increases and never achieve the l_0 bound with finite snapshot even in noiseless case. Recently, we showed that this drawback of the existing approaches can be overcome by the following generalized MUSIC criterion [7].

Theorem 1: [7] Assume that $A \in \mathbb{R}^{m \times n}$, $X \in \mathbb{R}^{n \times r}$, and $B \in \mathbb{R}^{m \times r}$ satisfy $AX = B$. Furthermore, we assume that $\|X\|_0 = k$ and A satisfies the RIP condition with the left RIP constant $0 < \delta_{2k-r+1}^L < 1$. If we are given $I_{k-r} \subset \mathrm{supp}X$ with $|I_{k-r}| = k - r$ and $A_{I_{k-r}} \in \mathbb{R}^{m \times (k-r)}$, which consists of columns whose indices are in I_{k-r}, then for any $j \in \{1, \cdots, n\} \setminus I_{k-r}$,

$$\mathbf{a}_j^* \left[P_{R(Q)} - P_{R(P_{R(Q)} A_{I_{k-r}})} \right] \mathbf{a}_j = 0 \tag{9}$$

if and only if $j \in \mathrm{supp}X$.

In [7], we demonstrate that the condition $0 < \delta_{2k-r+1}^L < 1$ for generalized MUSIC is equivalent to l_0 bound (6), which implies that a computational expensive combinatorial optimization problem is now reduced to $|I_{k-r}|$ support estimation from the original $|I_k|$ support estimation[1]. Furthermore,

---

[1]When $r = k$, the condition (9) is the same as the MUSIC criterion (7) and no combinatorial algorithm is necessary.

by Theorem 1, we can develop a computationally tractable relaxation algorithm called Compressive MUSIC (CS-MUSIC) that relaxes the combinatorial optimization step of finding I_{k-r} support using the conventional MMV-CS algorithms [7]. The algorithm can be stated as following:

- **(Step 1: compressed sensing step)** Find $k - r$ indices of $\text{supp}X$ by any MMV compressive sensing algorithms such as 2-thresholding or SOMP. Let I_{k-r} be set of selected indices and $S = I_{k-r}$.

- **(Step 2: generalized MUSIC step)** For $j \in \{1, \cdots, n\} \setminus I_{k-r}$, calculate the quantities $\eta(j) = \mathbf{a}_j^*[P_{R(Q)} - P_{R(P_{R(Q)}A_{I_{k-r}})}]\mathbf{a}_j$ for all $j \notin I_{k-r}$. Make an ascending ordering of $\eta(j)$, $j \notin I_{k-r}$ and choose indices that correspond to the first r elements and put these indices into S.

In compressive MUSIC, we determine $k - r$ indices of $\text{supp}X$ with CS-based algorithms such as 2-thresholding or S-OMP rather than l_0 optimization, where the exact identification of $k - r$ indices is a probabilistic matter. After that process, we recover remaining r indices of $\text{supp}X$ with a generalized MUSIC criterion, which is given in Theorem 1, and this reconstruction process is deterministic. This hybridization makes the compressive MUSIC applicable for all ranges of r, outperforming all the existing methods. Similar observation have been made independently by Lee and Bresler [17] in their subspace augmented MUSIC (SA-MUSIC) algorithm.

To analyze the performance of the compressive MUSIC, we should find the number of measurements with which we can identify the support of X. Due to the reduction of uncertainty from $|I_k|$ to $|I_{k-r}|$, we can expect more relaxed sampling condition. In [7], we derived the sampling requirements when subspace S-OMP or 2-thresholding is used as a compressed sensing step for compressive MUSIC. The results can be summarized as following. The number of measurements for subspace S-OMP for partial support recovery exhibits two distinct characteristics depending on the number of the measurement vectors. First, if the number of multiple measurement vectors r is sufficiently small, then the number of samples for S-OMP is reciprocally proportional to the number of multiple measurement vectors. On the other hand, we have sufficiently large number of snapshots such that $\lim_{n \to \infty}(\log n)/r$ is close to 0, then the number of measurements for S-OMP varies from $4k$ to k according to the ratio of r and k so that the $\log n$ is not necessary. In particular, if the number of snapshots approaches the sparsity k, then we can identify the indices of $\text{supp}X$ with only k measurements, which is equivalent to the required number of multiple measurement vectors for the success of conventional MUSIC. Furthermore, we demonstrated that the required

SNR for the success of support recovery can be reduced and when the asymptotic ratio of the number of snapshots and the sparsity level (that is, $\lim_{n\to\infty} r/k$) is nonzero in the large system limit, only finite SNR is required, which is significant improvement over SMV-CS.

In the original form of CS-MUSIC, the performance is, however, very dependent on the selection of $k - r$ correct indices of the support of X. In practice, even though the consecutive $k - r$ steps of S-OMP may not be correct, there are chances that among the estimates of k-sparse solution, part of the supports could be correct. Hence, if we have a mean to identify $k - r$ correct support in any order out of any k-sparse, then we can expect that the performance of the compressive MUSIC will be improved. Of course, when $\binom{k}{k-r}$ is small, we may apply the exhaustive search, but if both $k - r$ and r are not small, then the exhaustive search is hard to apply so that we have to find some alternative method to identify the correct indices from the estimate of $\mathrm{supp}X$. Indeed, the following support selection criterion can address the problem [8].

Theorem 2: [8] Assume that we have a canonical MMV model $AX = B$ where $A \in \mathbb{R}^{m \times n}$, $X \in \mathbb{R}^{n \times r}$, $\|X\|_0 = k$ and $r < k < m < n$. If there is an index set $I_k \subset \{1, \cdots, n\}$ such that $|I_k| = \min\{k, \mathrm{spark}(A) - r\}$ and $|I_k \cap \mathrm{supp}X| \geq k - r + 1$, then for any $j \in I_k$, $j \in \mathrm{supp}X$ if and only if

$$P_{Q_{k,j}} \mathbf{a}_j = \mathbf{0}, \tag{10}$$

where $Q_{k,j}$ is the orthogonal complement for $R([B \ A_{I_k \backslash \{j\}}])$, $A_{I_k \backslash \{j\}}$ consists of columns of A whose index belongs to $I_k \backslash \{j\}$ and $P^{\perp}_{R([B \ A_{I_k \backslash \{j\}}])}$ is the orthogonal projection on $R([B \ A_{I_k \backslash \{j\}}])^{\perp}$. In particular, if the columns of A are in general position, then we can take index set I_k with $|I_k| = \min\{k, m - r + 1\}$. Also, if A has an RIP condition with $0 < \delta_{2k} < 1$, then we can take $|I_k| = k$ since $r \leq k$.

Theorem 2 informs us that we only require the success of partial support recover out of k-sparse estimate, rather than $k - r$ consecutive correct CS step [7]. Accordingly, the compressive MUSIC with optimized partial support is then performed by following procedure.

- **[Step 1: compressed sensing]** Estimate k indices of $\mathrm{supp}X$ by any MMV compressive sensing algorithm. Let I_k be the set of indices which are taken in step 1.
- **[Step 2: support deletion]** For $j \in I_k$, calculate the quantities $\zeta(j) = \|P_{Q_{k,j}} \mathbf{a}_j\|^2$. Make an ascending ordering of $\zeta(j)$, $j \in I_k$ and choose indices that corresponds the first $k - r$ elements and put these indices into S and remove the remaining ones.

- **[Step 3: support addition]** For $j \in \{1, \cdots, n\} \setminus S$, calculate the quantities

$$\eta(j) = \mathbf{a}_j^* [P_{R(Q)} - P_{R(P_{R(Q)} A_{I_{k-r}})}] \mathbf{a}_j.$$

Make an asending ordering of $\eta(j)$, $j \notin S$ and choose indices that correspond to the first r elements and put these indices into S.

The step 1 in the above algorithm need not to be greedy so that we can also apply the convex optimization algorithm such as $l_{2,1}$ minimization [18] or belief propagation [19].

III. DETERMINISTIC SUPPORT TRACKING USING COMPRESSIVE MUSIC

A. Noiseless Cases

In this section, we will show how the compressive MUSIC with optimized partial support can be used for dynamic support tracking, whose joint support $\mathrm{supp}X(t)$ changes slowly along time as illustrated in Fig.1. First, we define a canonical form of dynamic MMV problem.

Definition 4: A canonical form of noiseless dynamic MMV problem is given by set of MMV problem with time varying k-sparse vectors $X(t) \in \mathbb{R}^{n \times r}$ that satisfies $Y(t) = AX(t)$ as described in following formulation:

$$\min_{X(t)} \|X(t)\|_0, \quad \text{subject to} \quad B(t) = AX(t), \quad t = 0, 1, \cdots, \tag{11}$$

where $\mathrm{supp}X(t) = \{1 \le i \le n : \mathbf{x}(t)^i \ne 0\}$ and $|\mathrm{supp}X(t)| = k(t)$, the measurement matrix $B(t)$ is full rank, i.e. $\mathrm{rank}(B(t)) \le k(t)$. Here we assume that $\mathrm{rank}(B(t))$ is constant so that we let $r := \mathrm{rank}(B(t))$.

Note that the canonical form MMV has the additional constraints that the measurement matrix is full rank and $\mathrm{rank}(B(t)) = r \le k(t)$. This is not problematic since every dynamic MMV problem can be converted into a canonical form using the following dimension reduction similar to [7] .

- Suppose we are given the following linear sensor observations: $B(t) = AX(t)$ where $A \in \mathbb{R}^{m \times n}$ and $X \in \mathbb{R}^{n \times l}$ satisfies $\|X(t)\|_0 = k(t)$.
- Compute the SVD as $B(t) = UD_r V^*$, where D_r is an $r \times r$ diagonal matrix, $V \in \mathbb{C}^{l \times r}$ consists of right singular vectors, and $r = \mathrm{rank}(B)$, respectively.
- Reduce the dimension as $B_{SV}(t) = B(t)V$ and $X_{SV}(t) = X(t)V$.
- The resulting canonical form MMV becomes $B_{SV}(t) = AX_{SV}(t)$.
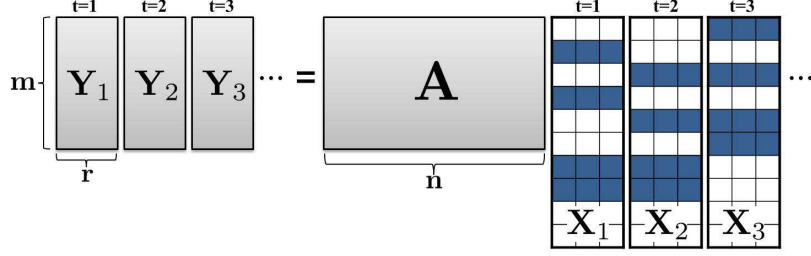
Fig. 1. MMV problem for slowly time varying sparsity pattern.

We can easily show that $\text{rank}(B_{SV}) = r \leq k(t)$ and full rank and the sparsity $k(t) := \|X(t)\|_0 = \|X_{SV}(t)\|_0$ with probability 1. Therefore, without loss of generality, the canonical form of dynamic MMV in Definition 4 is assumed throughout the paper.

For such dynamic support tracking, we can apply our CS-MUSIC algorithm. However, if the number of snapshots is not sufficient, the amount of support estimation that need to be done by CS step is significantly larger than those recovered by the deterministic generalized MUSIC step. Since CS step allows the support recovery in a probabilistic sense, it is more prone to error; so we are interested in finding a deterministic algorithm that significantly outperform the existing one. The following Theorem 3 shows that if we have a correct estimation for the initial support $I(0)$ of $X(0)$ and the support changes are sufficiently small and the sparsity $k(t)$ is fixed for all time point, then we can recursively identify the support of time-varying input signals in a *deterministic* manner.

Theorem 3: Suppose a noiseless canonical form of dynamic MMV problem satisfies

$$|\text{supp}X(t) \setminus \text{supp}X(t-1)| \leq r - 1, \tag{12}$$

for all $t = 1, 2, \cdots$. Furthermore we assume that $r \leq k(t) \leq k_{\max}$ for a positive integer $k_{\max}$ and $0 \leq \delta_{2k_{\max}}(A) < 1$. Then, if we have a correct initial support estimation for $X(0)$, then we can identify the correct support for all $t > 0$ by applying the following procedure recursively:

- **[Initial support estimation]** Let $I(t-1)$ be the support estimation of $X(t-1)$;
- **[Support deletion]** Find an index set $I(t)^a \subset I(t-1)$ such that $I(t)^a := \{j \in I(t-1) : \mathbf{a}_j^* P_{Q(t)_{k,j}} \mathbf{a}_j = 0\}$, where $Q(t)_{k,j}$ is the orthogonal complement for $R[B(t), A_{I(t-1)\setminus\{j\}}]$;

- **[Support addition]** Find an index set $I(t)$ such that

  $I(t) = \{j : \mathbf{a}_j^*[P_{R(Q(t))} - P_{R(P_{R(Q(t))}A_{I(t)^a})}]\mathbf{a}_j = 0\}$, where $Q(t) \in \mathbb{R}^{m \times (m-r)}$ consists of orthonormal columns such that $Q(t)^*B(t) = 0$;

- Set $\hat{k}(t) := |I(t)|$ be the sparsity estimate for $X(t)$ and $I(t)$ be the support estimate for $X(t)$.

  Proof: See Appendix A. ∎

In Theorem 3, we assume the RIP condition $0 \leq \delta_{2k_{\max}}^L(A) < 1$, instead of $0 \leq \delta_{2k_{\max}-r+1}^L(A) < 1$. If we assuming the RIP condition $0 \leq \delta_{2k_{\max}-r+1}^L(A) < 1$, when $r > 1 + k_{\max} - k(t)/2$, we may have $|I_k| < k(t)$. However, we can modify the support deletion procedure in Theorem 3 as the following, under the condition $|\text{supp}X(t) \setminus \text{supp}X(t-1)| \leq k_{\max}/2$.

- **[Support deletion]** Find an index set $I(t)^a \subset I(t-1)$ such that $I(t)^a := \{j \in I(t-1) : \mathbf{a}_j^* P_{Q(t)_{k,j}} \mathbf{a}_j = 0\}$, where $Q(t)_{k.j}$ is the orthogonal complement for $R[\tilde{B}(t), A_{I(t-1)\setminus\{j\}}]$ and $\tilde{B}(t)$ consists of $1 + [\frac{k_{\max}}{2}]$ columns of $B(t)$.

B. Noisy Cases

In practice, the measurements are noisy, so the theory we derived for noiseless measurement should be modified. In the noisy case, when the sparsity are known *a priori* and does not change along time, we can apply the following procedure.

- Let $t = 0$ and let $I(0)$ be the support estimation of $X(0)$.

- For all $t = 1, 2, \cdots$, do

  - Let $I(t) = \emptyset$.

  - For all $j \in I(t-1)$, calculate the quantities $\zeta(j) = \|P_{Q(t)_{j,k}} \mathbf{a}_j\|^2$.

  - Make an ascending ordering of $\zeta(j)$ and choose indices that correspond to the first $k - r$ elements and put these indices into $I(t)$.

  - For $j \in \{1, \cdots, n\} \setminus I(t)$, calculate the quantities $\eta(j) = \mathbf{a}_j^* \left[P_{R(Q(t))} - P_{R(P_{R(Q(t))}A_{I(t)})} \right] \mathbf{a}_j$.

  - Make an ascending ordering of $\eta(j)$, $j \notin I(t)$ and choose indices that correspond to the first r indices and add these indices to $I(t)$.

  - $I(t)$ is the estimation of $\text{supp}X(t)$ and let $t = t + 1$.

However, if the sparsity changes along time, in the noisy cases, some of the steps in Theorem 3 should be modified as follows:

- **[Support deletion]** Set $\epsilon_1 > 0$ and find an index set $I(t)^a$ such that $I(t)^a = \{j \in I(t-1):$
  $\mathbf{a}_j^* P_{Q(t)_{k,j}} \mathbf{a}_j < \epsilon_1\}$ where $Q(t)_{k,j}$ is the orthogonal complement for $R[Y(t)\ A_{I_1(t)\backslash\{j\}}]$, where
  $I_1(t) \subset I(t-1)$ such that

$$\mathrm{nrank}[Y(t)\ A_{I_1(t)}] = \mathrm{nrank}[Y(t)\ A_{I(t-1)}] = r + |I_1(t)|,$$

  where $\mathrm{nrank}(A)$ denotes the numerical rank of A.

- **[Support addition]** Set $\epsilon_2 > 0$ and find an index set $I(t)^b$ such that

$$I(t)^b = \{j \notin I(t)^a : \mathbf{a}_j^* P_{R([Y(t)\ A_{I_2(t)}])^\perp} \mathbf{a}_j < \epsilon_2\},$$

  where an index set $I_2(t) \subset I(t)^a$ such that

$$\mathrm{nrank}[Y(t)\ A_{I_2(t)}] = \mathrm{nrank}[Y(t)\ A_{I(t)^a}] = r + |I_2(t)|.$$

In this section, we derive sufficient conditions for the threshold values and signal to noise ratio that guarantee the correct identification of time varying support. For CS-MUSIC [7], we derived an expression of SNR and the minimum number of sensor elements. Even though these derivation is based on a large system model with a Gaussian sensing matrix, it has provided very useful insight. Therefore, we employed a large system model to derive a sufficient condition for the success of proposed algorithm.

Definition 5: A large system noisy canonical form of dynamic MMV is defined as an estimation problem of $k(t)$-sparse vectors $X(t) \in \mathbb{R}^{n \times r}$ that shares a common sparsity pattern through multiple noisy snapshots $Y(t) = AX(t) + N(t)$ using the following formulation:

$$\mathrm{minimize} \quad \|X(t)\|_0 \tag{13}$$

$$\mathrm{subject\ to} \quad Y(t) = AX(t) + N(t),$$

where $A \in \mathbb{R}^{m \times n}$ is a random matrix with i.i.d. $\mathcal{N}(0, 1/m)$ entries, $N = [\mathbf{n}_1, \cdots, \mathbf{n}_r] \in \mathbb{R}^{m \times r}$ is an additive noise matrix, $m \to \infty$, $k \to \infty$ as $n \to \infty$ and $\mathrm{rank}(AX(t)) = r(t) \le k(t) = \|X(t)\|_0$. Here, we assume that $\rho := \lim_{n \to \infty} m/n > 0$ and $\gamma = \lim_{n \to \infty} k_{\max}/m > 0$, $\alpha := \lim_{n \to \infty} r/k_{\max} \ge 0$ exist and $\alpha \le 1 - \epsilon$ for some $0 < \epsilon < 1$.

Under the large system model, we have the following theorem.

Theorem 4: Consider the large system model dynamic MMV in Definition 5. Suppose a minimum SNR satisfies

$$\mathsf{SNR}_{\min}(Y(t)) := \frac{\sigma_{\min}(B(t))}{\|N\|} > 1 + \frac{4(\kappa(B(t)) + 1)}{1 - \gamma(1 + \alpha)}, \tag{14}$$

where $\sigma_{\min}(B(t))$ is the minimum singular value for $B(t)$, $\|N\|$ is the spectral norm of $N \in \mathbb{R}^{m \times r}$ and $B(t)$ is the noiseless measurements, and $\alpha = \lim_{n \to \infty} r/k_{\max}$, $\gamma = \lim_{n \to \infty} k_{\max}/m$. Then, for the noisy canonical form dynamic MMV problem for slowly time varying pattern that satisfies Eq. (12), the threshold values for support deletion and addition criterion to the correct partial support for $X(t)$ are given by

$$\epsilon_1 := (1 - \gamma(1 + \alpha))/2, \qquad \epsilon_2 := (1 - \gamma)/2. \tag{15}$$

Proof: See Appendix B. ∎

IV. NUMERICAL RESULTS

The first simulation is to demonstrate the performance of the proposed method to solve the time varying MMV problem in Eq. (11) for different number of changes in supports at each time. We declared the algorithm as a success if the estimated support is the same as the true $\mathrm{supp}X$, and the success rates were averaged for 5000 experiments. The simulation parameters were as follows: $m = 40$, $n = 100$, $r = 9$, and $k \in \{1, 2, \cdots, 30\}$, respectively. Elements of sensing matrix A were generated by i.i.d. Gaussian random variable $\frac{1}{\sqrt{m}} \mathcal{N}(0, 1)$, and Gaussian noise of $\mathrm{SNR} = 40 dB$ was added to each measurement vectors. At each time point, $X(t)^{\mathrm{supp}X(t)}$ is generated by $\mathcal{N}(0, 1)$. Fig.2 shows the recovery rates of time varying MMV problem using support tracking method for $t = 1, 2, \cdots, 5$ when the number of changed supports are $4, 6, 7$, and 8 at each time point for Fig.2(a)$\sim$(d), respectively. We used CS-MUSIC algorithm with S-OMP and then applied optimized partial support selection at $t = 1$, and time varying supports are estimated by support tracking method recursively from $t = 2$ to $t = 5$. In Fig.2, we can observe that the performance gracefully decreases as the number of changes in supports increases. An interesting observation is that the performance of the proposed method rather improves over time in Fig.2(a) and (b). However, the recovery ratio is getting lower but converges over time when the number of changes in supports is close to the upper bound $r - 1$ for perfect recovery in noiseless case.

Next, we applied the proposed algorithm to target tracking problem in $2D$ image and compared it to MUSIC algorithm. The first row of Fig.3 indicates the original targets moving toward the direction of red arrows over time. Each column (from left to right) indicates the sampled image at $t = 1, 13, 27$, and $t = 41$, respectively. The simulation setting is the same with the previous one except $m = 50$, $n = 900$, $t = 1, 2, \cdots, 45$, and each target have a chance to move with
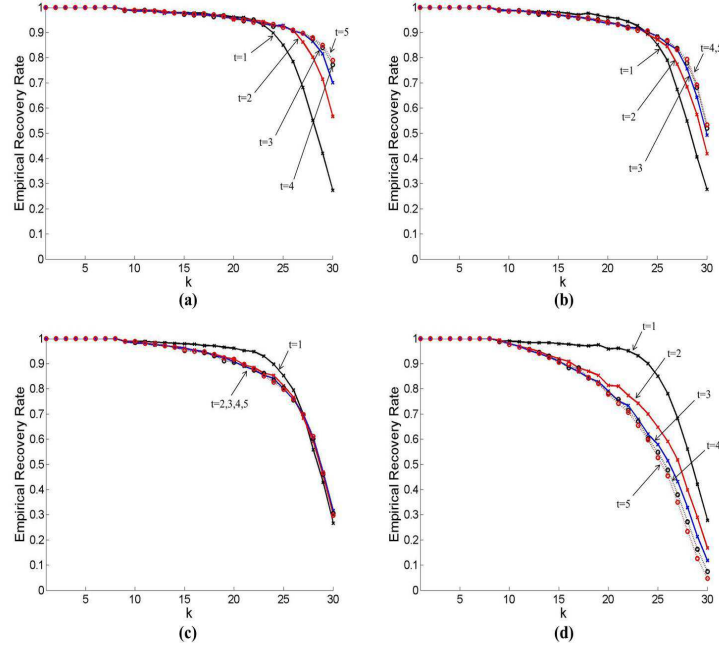
Fig. 2. Recovery rates of time varying MMV problem using support tracking method when $m = 40$, $n = 100$, $r = 9$, SNR= 40dB, and $t = 1, 2, \cdots, 5$. The number of changes in supports at each time point is (a) 4, (b) 6, (c) 7, and (d) 8.

probability of $\frac{1}{2}\frac{r-1}{k}$ at each time point. The number of target k is 24. Here, we considered the number of measurement vectors is 50 in the resting state, and used MUSIC algorithm to find supports at $t = 0$. The second and third row of Fig.3 indicate the results of support tracking method and MUSIC algorithm, respectively. Note that the proposed method successfully follows the movement of original targets.

V. CONCLUSION

This paper expanded the sparse recovery with partially known supports in single measurement vector problem to multiple measurement vector problem with joint sparsity and proposed the support tracking algorithm to recover the slowly time varying supports. It is based on the recently developed compressive MUSIC algorithm with optimized partial support selection. The estimated supports at previous time can be used in optimized partial support selection to recover partial supports at current time and it can be used in generalized MUSIC criterion to find remaining supports. We also provided the maximum allowable number of changes in supports with support
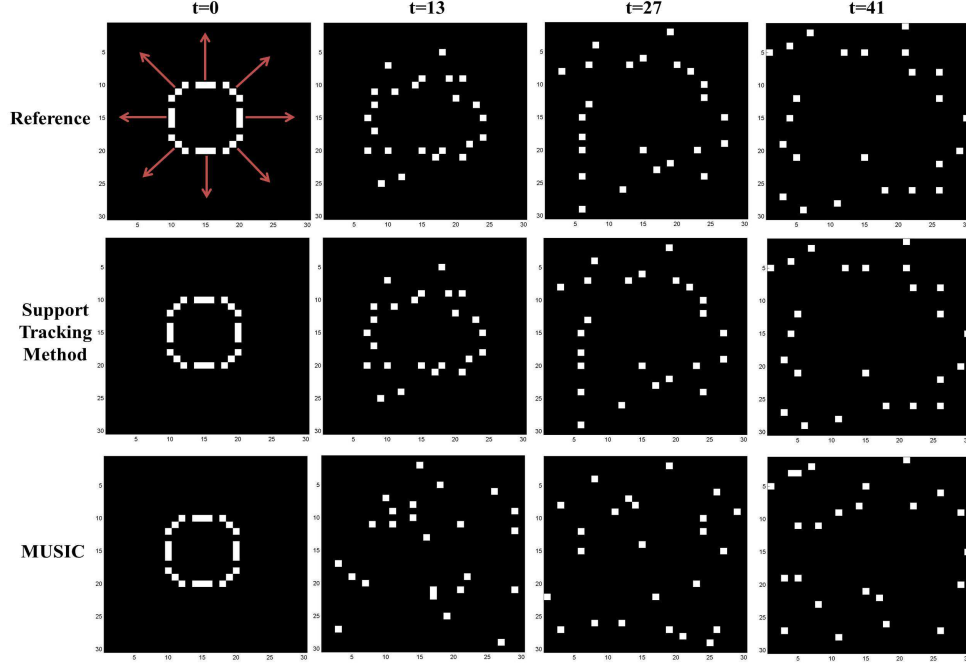
Fig. 3. The results of the target tracking problem in $2D$ image when $m = 50$, $n = 900$, $k = 24$, and SNR= 40dB. We set $r = 50$ when $t = 0$, and $r = 9$ for $t > 0$. The first row indicates the original targets moving toward the direction of red arrows over time. The second and third row indicate the results of support tracking method and MUSIC algorithm, respectively. Each column (from left to right) indicates the sampled image at $t = 1, 13, 27$, and $t = 41$, respectively.

tracking algorithm for exact reconstruction in noiseless case. Numerical results demonstrated that the proposed algorithm reliably reconstructs the time varying supports for various level of changes and successfully solves the target tracking problem in $2D$ image.

APPENDIX A

Proof: We only need to show that if we have a correct support for $X(t-1)$, then we can also obtain a correct support estimation for $X(t)$ by the support selection criterion and the generalized MUSIC criterion. By the assumption, we have $m \geq 2k_{\max} \geq k(t-1) + r$ so that if we have $|\text{supp}X(t) \cap I(t-1)| \geq k(t) - r + 1$, then by Theorem 2 we have for any $j \in I(t-1)$,

$$j \in \text{supp}X(t) \text{ if and only if } \mathbf{a}_j^* P_{Q(t)_{k,j}} \mathbf{a}_j = 0$$

where $Q(t)_{k,j}$ is the orthogonal complement of $R([Y(t) \ A_{I(t-1)\setminus\{j\}}])$. Since we have a noiseless MMV problem with slowly time varying pattern, we have $|\text{supp}X(t) \setminus \text{supp}X(t-1)| \leq r - 1$ so

that we have $|\mathrm{supp}X(t) \cap I(t-1)| \geq k(t) - r + 1$ and we can identify the correct partial support of $X(t)$ which has at least $k(t) - r + 1$ elements. Then, if we let $I(t)^a$ be the set of indices such that

$$I(t)^a = \{j \in I(t-1) : \mathbf{a}_j^* P_{Q(t)_{k,j}} \mathbf{a}_j = 0\},$$

we have $I(t)^a \subset \mathrm{supp}X$ and $R([Y(t)\ A_{I(t)^a}]) \subset R(A_{\mathrm{supp}X(t)})$. On the other hand, if we take a set $I(t,r) \subset I(t)^a \subset \mathrm{supp}X$ such that $|I(t,r)| = k(t) - r$, we have

$$R([Y(t)\ A_{I(t)^a}]) \supset R([Y(t)\ A_{I(t,r)}]) = R(A_{\mathrm{supp}X(t)})$$

which implies $R([Y(t)\ A_{I(t)^a}) = R(Y(t)\ A_{I(t,r)})$. Since $0 \leq \delta_{2k(t)-r+1}(A) \leq \delta_{2k_{\max}-r+1}(A) < 1$, we can apply the generalized MUSIC criterion with $I(t,r) \subset \mathrm{supp}X$ where $|I(t,r)| = k(t) - r$. For $j \in I(t,r)$, we can easily see that

$$\mathbf{a}_j^* \left[P_{R(Q(t))} - P_{R(P_{R(Q(t))}A_{I(t)^a})} \right] \mathbf{a}_j = \mathbf{a}_j^* P_{R([Y\ A_{I(t)^a}])^\perp} \mathbf{a}_j = \mathbf{a}_j^* P_{R([Y\ A_{I(t,r)}])^\perp} \mathbf{a}_j = 0.$$

On the other hand, for $j \notin I(t,r)$, by the generalized MUSIC criterion, we have $j \in \mathrm{supp}X(t)$ if and only if

$$\mathbf{a}_j^* \left[P_{R(Q(t))} - P_{R(P_{R(Q(t))}A_{I(t)^a})} \right] \mathbf{a}_j = \mathbf{a}_j^* \left[P_{R(Q(t))} - P_{R(P_{R(Q(t))}A_{I(t,r)})} \right] \mathbf{a}_j = 0.$$

Since $I(t,r) \subset \mathrm{supp}X$, we have $j \in \mathrm{supp}X$ if and only if

$$\mathbf{a}_j^* \left[P_{R(Q(t))} - P_{R(P_{R(Q(t))}A_{I(t)^a})} \right] \mathbf{a}_j = 0.$$

Hence, $|I(t)| = k(t)$ and $I(t) = \mathrm{supp}X(t)$. ∎

APPENDIX B

Proof: Here, we let $B(t) = AX(t)$, $\sigma_{\min}(B(t))$(or $\sigma_{\min}(B(t))$) be the minimum (or the maximum) nonzero singular value of $B(t)$. Then $Y(t) = B(t) + N(t)$ is also of full column rank if $\|N(t)\| < \sigma_{\min}(B(t))$. By [7], for such an $N(t)$, we have

$$\|P_{R(Y(t))} - P_{R(B(t))}\| \leq \frac{2[\sigma_{\max}(B(t)) + \sigma_{\min}(B(t))]\|N(t)\|}{\sigma_{\min}(B(t))(\sigma_{\min}(B(t)) - \|N(t)\|)}. \tag{16}$$

By the projection update rule, we have

$$P_{R([B(t)\ A_{I_1(t)\setminus\{j\}}])} = P_{R(A_{I_1(t)\setminus\{j\}})} + P_{R(P_{R(A_{I_1(t)\setminus\{j\}})}^\perp B(t))} \tag{17}$$

and

$$P_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])} = P_{R(A_{I_1(t)\setminus\{j\}})} + P_{R(P^\perp_{R(A_{I_1(t)\setminus\{j\}})}Y(t))}. \tag{18}$$

Since $[B(t)\ A_{I_1(t)\setminus\{j\}}]$ and $[Y(t)\ A_{I_1(t)\setminus\{j\}}]$ are of full column rank, by applying (17) and (18) as done in [17], we have

$$\begin{aligned}
\|P^\perp_{R([B(t)\ A_{I_1(t)\setminus\{j\}}])} - P^\perp_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])}\| &= \|P_{R(P^\perp_{R(A_{I_1(t)\setminus\{j\}})}B(t))} - P_{R(P^\perp_{R(A_{I_1(t)\setminus\{j\}})}Y(t))}\| \\
&\leq \|P_{R(B(t))} - P_{R(Y(t))}\|.
\end{aligned} \tag{19}$$

Then for any $j \in I_1(t) \setminus \mathrm{supp}X$, we have

$$\begin{aligned}
\mathbf{a}_j^* P^\perp_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])}\mathbf{a}_j &= \mathbf{a}_j^* P^\perp_{R([B(t)\ A_{I_1(t)\setminus\{j\}}])}\mathbf{a}_j \\
&\quad + \mathbf{a}_j^* \left[P^\perp_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])} - P^\perp_{R([B(t)\ A_{I_1(t)\setminus\{j\}}])} \right] \mathbf{a}_j \tag{20} \\
&\geq \min_{j\notin\mathrm{supp}X} \mathbf{a}_j^* P^\perp_{R([B(t)\ A_{I_1(t)\setminus\{j\}}])}\mathbf{a}_j - \max_{1\leq j\leq n} \|\mathbf{a}_j\|^2 \|P_{R(Y(t))} - P_{R(B(t))}\|.
\end{aligned}$$

Here, for each $1 \leq j \leq n$, $m\|\mathbf{a}_j\|^2$ is a chi-square random variable with degree of freedom m so that we have by Lemma 3 in [20], $\lim_{n\to\infty} \max_{1\leq j\leq n} \|\mathbf{a}_j\|^2 = 1$ since $\lim_n (\log n)/m = 0$. Furthermore, for any $j \notin \mathrm{supp}X$, $\mathbf{a}_j$ is independent of $P^\perp_{R([Y(t)\ A_{I_{(t-1)}\setminus\{j\}}])}$, so that $m\mathbf{a}_j^* P^\perp_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])}\mathbf{a}_j$ is a chi-squared random variable whose degree of freedom is at least $m - k(t) - r + 1$ since $P^\perp_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])}$ is a projection operator onto the orthogonal couplement of $R([Y(t)\ A_{I_1(t)\setminus\{j\}}])$. Since $\lim_{n\to\infty}(\log(n - k(t)))/(m - k(t) - r + 1) = 0$, again by Lemma 3 in [20], we have

$$\lim_{n\to\infty} \frac{\min\limits_{j\notin\mathrm{supp}X} m\mathbf{a}_j^* P^\perp_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])}\mathbf{a}_j}{m - k(t) - r + 1} \geq 1$$

so that

$$\lim_{n\to\infty} \min_{j\notin\mathrm{supp}X} \mathbf{a}_j^* P^\perp_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])}\mathbf{a}_j \geq 1 - \gamma(1 + \alpha) \tag{21}$$

since $k(t) \leq k_{\max}$ for all $t = 0, 1, \cdots$. On the other hand, if we use the definition of $\mathsf{SNR}_{\min}(Y(t))$ and the definition of the condition number of $B(t)$ on (16), i.e. $\kappa(B(t)) = (\sigma_{\max}(B(t)))/(\sigma_{\min}(B(t)))$, we have

$$\|P_{R(Y(t))} - P_{R(B(t))}\| \leq \frac{2(\kappa(B(t)) + 1)}{\mathsf{SNR}_{\min}(Y(t)) - 1} < \frac{1 - \gamma(1 + \alpha)}{2}, \tag{22}$$

by the condition (14). Combining (20), (21) and (22), we have for any $j \in I(t - 1) \setminus \mathrm{supp}X$, we have

$$\mathbf{a}_j^* P^\perp_{R([Y(t)\ A_{I_1(t)\setminus\{j\}}])}\mathbf{a}_j > \frac{1 - \gamma(1 + \alpha)}{2}.$$

On the other hand, for $j \in I(t-1) \cap \mathrm{supp} X(t)$, we have $\mathbf{a}_j^* P_{R([B(t) \ A_{I_1(t) \setminus \{j\}}])}^{\perp} \mathbf{a}_j = 0$ by the support selection criterion. Then, by the similar reasoning as above, we have for any $j \in I_1(t) \cap \mathrm{supp} X(t)$, we have

$$\mathbf{a}_j^* P_{R([Y(t) \ A_{I_1(t) \setminus \{j\}}])}^{\perp} \mathbf{a}_j < \frac{1 - \gamma(1 + \alpha)}{2}.$$

This completes the proof for the threshold values for support selection criterion. The proof for the threshold values for generalized MUSIC are the same except that $m\mathbf{a}_j^* P_{R([Y(t) \ A_{I_2(t)}])}^{\perp} \mathbf{a}_j$ is a chi-squared random variable whose degree of freedom is $m - k(t)$ for $j \notin \mathrm{supp} X(t)$. ∎

REFERENCES

[1] Z. Zhang and B.D. Rao, "Sparse signal recovery with temporally correlated source vectors using joint sparse bayesian learning," *IEEE J. of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912–926, 2011.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[3] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[4] N. Vaswani and W. Lu, "Modified-CS: modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4595–4607, 2010.

[5] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[6] Y. Jin and B.D. Rao, "Support recovery of sparse signals in the presence of multiple measurement vectors," *arXiv preprint*, 2011, http://arxiv.org/pdf/1109.1895.

[7] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive MUSIC: a missing link between compressive sensing and array signal processing," *to appear in IEEE Trans. Inf. Theory*, 2011.

[8] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive MUSIC with optimized partial support for joint sparse recovery," *in Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2011.

[9] D. L. Donoho, "Neighborly polytopes and sparse solution of underdetermined linear equations," *Tech. report, Department of Statistics, Stanford University*, 2005.

[10] D. L. Donoho and M. Elad, "Optimally sparse representation in general (non-orthogonal) dictionaries via l_1 minimization," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 5, pp. 2197–2202, 2003.

[11] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple measurement vectors," *IEEE Trans. on Signal Processing*, vol. 54, no. 12, pp. 4634–4643, 2006.

[12] P. Feng, *Universal minimum-rate sampling and spectrum-blind reconstruction for multiband signals*, Ph.D. dissertation, University of Illinois, Urbana-Champaign, 1997.

[13] M. E. Davies and Y. C. Eldar, "Rank awareness for joint sparse recovery," *preprint*, 2010, http://arxiv.org/PS_cache/arxiv/pdf/1004/1004.4529v1.pdf.

[14] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[15] R. Gribonval, H. Rauhut, K. Schnass, and P. Vandergheynst, "Atoms of all channels, unite! Average case analysis of multi-channel sparse recovery using greedy algorithms," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 655–687, 2008.

[16] Y.C. Eldar and H. Rauhut, "Average case anlysis of multichannel sparse recovery using convex relaxation," *IEEE Trans. on Information Theory*, vol. 56, pp. 505–519, 2010.

[17] K. Lee and Y. Bresler, "Subspace-augmented music for joint sparse recovery," *preprint*, 2010, http://arxiv.org/PS_cache/arxiv/pdf/1004/1004.3071.pdf.

[18] D. Malioutov, M. Cetin, and AS Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[19] J.M. Kim, W.H. Chang, B.C. Jung, D. Baron, and J.C. Ye, "Belief propagation for joint sparse recovery," *arXiv preprint*, 2011, http://arxiv.org/pdf/1102.3289.

[20] S. Rangan A.K. Fletcher and V.K. Goyal, "Necessary and sufficient conditions for sparsity pattern recovery," *IEEE Trans. on Inform. Theory*, vol. 55, no. 12, pp. 5758–5772, December 2009.